DOCUMENT RESUME

ED 185 072 TM 800 058

AUTHOR Richardson, Ellis: And Others
TITLE The Use of The Decoding Skills Test in the Measurement of Auditory and Visual Work Recognition Skills.
PUB DATE [Apr 79]
NOTE 44p.: Paper presented at the Annual Convention of the International Reading Association (24th, Atlanta, GA, April 23-27, 1979).

EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS *Decoding (Reading): Elementary Education: Oral Reading: *Phonics: Reading Achievement: Reading Processes: *Reading Tests: *Sight Vocabulary: Test Construction: Test Reliability: *Test Validity: *Word Recognition
IDENTIFIERS *Decoding Skills Test: Passage Error Rate: Phonic Transfer Index

ABSTRACT

A model for the measurement of reading skills was used as the basis for the development of the Decoding Skills Test. The model emphasizes that decoding skills are based on one of two processes: word recognition through phonic/linguistic decoding or through basal word recognition. Word lists appropriate for the elementary grades were used to construct the Basal Word Recognition subtest. The Phonic Decoding subtest was based on Venezky's common spelling patterns, and the concepts associated with syllables and meaning. The Oral Reading subtest presented words in context which had already been seen (out of context) on the previous two subtests. The test was administered to approximately 600 elementary school students; other standardized reading achievement tests were administered to select ten good and ten poor readers from each grade level. Test validity, discriminating ability, and interrater and internal consistency reliability were examined. (Findings are discussed, and several test items are appended). (GDC)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

Form 91.28 Adm

State of New York
Department of Mental Hygiene
LONG ISLAND RESEARCH INSTITUTE
State University of New York at Stony Brook
Health Sciences Center 10T
Stony Brook, New York 11794

STANLEY F. YOLLES, M.D., M.P.H., F.A.C.P.
INSTITUTE DIRECTOR
(516) 444-2163

DIV. OF CHILD MENTAL HEALTH
BERTRAND WINSBERG, M.D.
DIRECTOR

U S DEPARTMENT OF HEALTH,
EDUCATION & WELFARE
NATIONAL INSTITUTE OF
EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

ED185072

The Use of the Decoding Skills Test in the Measurement of Auditory and Visual Word Recognition Skills

Ellis Richardson, Barbara DiBenedetto
Long Island Research Insititue
and
Arlene Adler
Downstate Medical Center, SUNY

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY
Ellis Richardson
TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

Paper presented at the Twenty-Fourth Annual Convention of the International Reading Association

TM800 058

The Use of the Decoding Skills Test in the Measurement of Visual and Auditory Word Recognition Skills

One of the most important theoretical assumptions in the field of reading has been summarized by Vellutino and his associates: "...word identification (is) a complex process which involves the retrieval of information from both (visual and linguistic) storage systems." (Vellutino, Smith, Steger, & Kaman, 1975, p. 493). Most reading programs can be divided into meaning emphasis (visual) and code-emphasis (auditory) approaches (see Chall, 1967). A number of instruments have been devised which purport to measure skills in these domains (e.g., Frostig, 1966; Goldman, Fristoe & Woodcock, 1974; Kirk, McCarthy & Kirk, 1968), and these have been used to study auditory and visual skill correlates of reading (e.g., Golden & Steiner, 1969, Olson, 1966) and to study differences between good and poor readers (e.g., Bruininks, Lucker & Gropper, 1970; Macione, 1969). However, little usefull information has been gained from these efforts (see Cohen, 1969; Hammill & Larsen, 1974; Larsen, Rogers, & Sowell, 1976). Nevertheless, the compelling fact remains: code emphasis programs rely upon the use of abstract sound-symbol relationships to produce potentially meaningless utterances (i.e., appear to stress auditory systems) while meaning-emphasis approaches rely upon the association of abstract visual images with only meaningful utterances (i.e., appear to stress visual systems).

While much effort has been invested in the search for the auditory and visual correlates of reading, very little has been done to study the acquisition of actual reading skills which should

be related to these two domains. This may be, in part, because adequately developed instruments to make such comparisons have not been available. This paper describes a model for the measurement of phonic (auditory) and sight-word (visual) skills as they are integrated in the process of reading. The development of the Decoding Skills Test (DST) which is based on the model is reviewed, and a study in which the test was used to compare good and poor readers is then presented.

The DST was developed by the Long Island Research Institute through a contract with the National Institute of Child Health and Human Development. The initial purpose in developing the DST was to provide a "detailed measurement of decoding skills which could be used in research on developmental dyslexia." (RFP #77-18, 1977). The DST is criterion-referenced in that its contents have been derived from a wide variety of classroom materials currently used in Grades One through Five. It consists of three subtests: Basal Word Recognition (Subtest I), Phonic Decoding (Subtest II), and Oral Reading (Subtest III).

The Model

The DST is based on the simplified model of decoding in reading shown in Figure 1. In this model, primary emphasis is placed on word recognition. The model posits two ways in which a child may recognize a word--through phonic/linguistic decoding or through basal word recognition--which are then combined in contextual reading. Although the two methods of word recognition are not mutually exclusive (since basal vocabulary may be based,

at least in part, on words containing phenetically "regular" spelling patterns), for purposes of measurement, they are posed as seperate skill areas in the DST.

Basal vocabulary. A child may be able to decode a word because it has been presented in classroom lessons or because he/she has been exposed to the word through other reading experiences. Certain words receive high exposure during classroom instruction and appear with high frequency in most basal reader programs. Although many such basal vocabulary words are phonetically regular (in that their pronunciations are governed by common spelling-to-sound correspondence rules), many are irregular, and others are regular, but contain complex spelling-to-sound correspondence patterns.

Phonic/linguistic decoding. A child may be able to decode a word because he/she has learned the sounds that most commonly correspond to the particular letters (or groups of letters) in the word. Most words in the language do conform to common spelling-to-sound correspondence patterns. Some of these patterns are simple (e.g., short vowel trigrams; long vowel, silent *e* ending words) and others are more complex (e.g., words containing vowel digraphs). A child will be able to apply knowledge of spelling-to-sound correspondence in decoding phonetically regular words only to the degree that particular patterns have been mastered.

Oral reading. Both methods of word recognition are combined in the act of contextual reading where they interact with a child's ability to use context clues and his/her overall fluency in reading.

The Subtests: Design and Development

Subtest I: Basal Word Recognition

The primary objective of the Basal Word Recognition Subtest is to provide a criterion-referenced estimate of a child's ability to read words at various instruction levels in Grades One through Five. Most primary level reading series may be divided into three reader levels in the First Grade (preprimer, primer and first reader) and two reader levels in each grade thereafter (e.g., 2-1 and 2-2; 3-1 and 3-2; etc.) The subtest is composed of eleven lists of 15 words each, which were selected to be representative of the instructional content of most reading programs at each of the eleven reader levels in Grades One through Five.[1] Thus, a child's mastery of the vocabulary at a given reader level is estimated from his/her score on each of the reader-level lists.

Analysis of basal reader vocabularies. The procedure for assembling words for the Basal Word Recognition Subtest involved analyzing the vocabularies of a number of widely used reading programs to find the words which had the highest likelihood of being taught at each of the eleven reader levels (This likelihood was defined by the number of instructional programs which presented a given word at a given level.) The ten different reading programs used in this analysis are shown in Table 1. An effort was made to sample recently published programs (Houghton-Mifflin, MacMillan "r", and Book Mark), programs which have been in wide use for a number of years (Allyn-Bacon, Bank Street, and MacMillan), and programs representing linguistic and phonic (i.e., code emphasis) approaches to vocabulary control (SRA/BRS, Merrill, Programmed Reading, Lippincott).

The 15 most frequently taught words at each reader level were then selected for Subtest I. Table 2 shows the words selected for the 2-1 reader list, together with relevant information used in the selection process. The first column in Table 2 shows the actual words selected. The next two columns show the number of meaning-emphasis (Mean.) and code-emphasis (Code) programs which introduced that word at the given reader level (i.e., 2-1). The third column shows the number of programs (including both meaning-emphasis and code-emphasis) which introduced the word at a reader level prior to the given one, and the fourth column shows the cumulative number of programs which had introduced the word at or by the given level. Thus, the word "better" was introduced at the 2-1 level by five of the meaning-emphasis programs and by two of the code-emphasis programs. It had been introduced in the First Grade material of two of the remaining three programs and, thus, had been presented by a total of 9 of the 10 programs surveyed by the 2-1 reader level. The last two columns in Table 2 shows shows the cross validation of these words with the Botel (1962) and Harris & Jacobson (1972) word lists. For example, the word "better" is listed at the 2-1 reader level by Botel, and as a First Grade word by Harris & Jacobson. It will be noted that the Botel ratings tended to be a bit higher than the DST ratings, while those of Harris & Jacobson tended to be a bit lower. The overall correspondence between these lists and the DST items is, however, remarkably close. These relationships tended to hold for all reader level lists.

Response recording and administration procedures. Correct responses to each item are classified by response latency estimates: (a) those which occur immediately (within an estimated 2 seconds), (b) those which require more than 2 but less than 5 seconds, and (c) those which require 5 or more seconds. Incorrect responses are also classified into three categories: (a) no response, (b) incorrect responses which are based on letter-sound clues in the word (e.g., "stop" for "shop"), and (c) incorrect responses which are totally different from the correct responses (defined as having fewer than two sounds in common with the correct response [e.g., "stay" for "shop"]).

When the test is administered, a child is started on the reader list which is one full grade level below the child's grade placement. For example, a child in the latter half of the Third Grade would be started on the 2-2 list; a child in the first half of the Fifth Grade would be started on the 4-1 list. If the child misses more than 80% of the items on that list, the previous list is given until the child scores at least 80% of the items correct. Successively higher lists are then given until the child scores less than 50% of the items on a particular list correct (or completes the 5-2 list) and the subtest is terminated. This procedure permits the establishment of a "basal" reader level--the highest list on which the child scores at least 80% correct, and a "ceiling" reader level--the lowest list on which the child is unable to read at least 50% of the words.

Subtest II: Phonic Decoding

The primary objective of the Phonic Decoding Subtest is to provide an estimate of a child's ability to recognize or pronounce words on

the basis of spelling patterns. The most common spelling-to-sound relationships are described by the major patterns analyzed by Venezky (1970). The major patterns consist of the most common vowel-sound spellings which we have divided into three categories: (a) the short vowel, generally signified by the short vowel trigram, CVC (e.g., "cat," "pot"), (b) the long vowels, generally signified by the silent *e* endings, CVC*e* (e.g., "time," "same"), and (c) the diphthongs and vowel digraphs, CVVC (e.g., "boat," "boil," "dawn"). A second dimension to the spelling patterns may be introduced by making a distinction between words with only single consonants and those with consonant blends and digraphs.

The Phonic Decoding Subtest is designed to sample words from each of the six spelling patterns defined by the three vowel types and two consonant types. Two additional dimensions are added to the design of the subtest: syllables and meaning. The subtest includes examples of monosyllabic words representing each of the six patterns, as well as examples of polysyllabic words also representing the same pattern, making a total of 12 cells. For each of these 12 cells, the subtest includes examples of real (i.e., meaningful) words and nonsense words, making a total of 24 cells. Five items are included in each of the 24 cells, for a total of 120 words on the Phonic Decoding Subtest.[2]

Development of the Phonic Decoding Subtest. The fundamental assumption of the Phonic Decoding Subtest is that the child reads a given "real" word on the basis of a knowledge of the pronunciation of the spelling pattern rather than on the basis of classroom instruction. On the other hand, the assumption in the testing of the real words is that the word is in the child's speaking vocabulary (i.e., is a real word for the child). Thus, an effort was made to find words with a

minimum likelihood of having been included in elementary school reading programs and a maximum likelihood of being in the speaking vocabulary of most elementary school children. The final item selection resented a compromise between these two considerations. The Harris & Jacobson (1972) list was used as a source of words, to provide items with a high likelihood of being in the children's speaking vocabularies. However, only words which did not appear in the DST analysis of basal vocabularies as frequently taught words were selected, thus minimizing the probability that the words would have been presented in more current classroom materials.

The 60 words selected for the Real Word section of Subtest II are shown in Table 3. The monosyllabic words were selected to represent First through Fifth Grade levels.[3] Thus, "cat" is listed as a First Grade word by Harris & Jacobson, the word "hit" is listed as a Second Grade word, and so on. Similarly, "drop" is listed as a First Grade word, "shut"is a Second Grade word, "path," a Third Grade word, etc. Since there are no examples of First Grade polysyllabic words which conform to several of the spelling patterns, the first polysyllabic word in each cell represents the Second Grade level (e.g., "visit," "finish," "beside"). Succeeding words in each cell are then drawn from successive grade levels, so that the last word in each list (e.g., "atomic," "stiffen," "awoke") represents Harris & Jacobson's Sixth Grade level.

After the 60 Real Words had been selected for the Phonic Decoding Subtest, the 60 Nonsense words were formed simply by changing the initial consonant sound of each Real Word. For example, "cat" became "lat," drop" became "brop," and so on (see Table 3).

Response recording and administration procedures. The same categories used for classifying correct and incorrect responses on the Basal Word Recognition Subtest were employed on the Phonic Decoding Subtest, with the exception that an additional incorrect response category was used. Incorrect responses which followed an obvious attempt on the child's part to sound out the word were classified seperately. When the test is administered, the child is told that the letter sounds may aid him/her in determining what the word is. The five words in each cell are then exposed in succession beginning with the monosyllabic Real Words: CVC words, followed by CCVVCC, CVCe, CCVCe, CVVC, and finally CCVVCC words. The subtest is continued with the exposure of the polysyllabic Real Words in the same sequence. If or when the child misses 10 consecutive words, the testing of the Real Words is terminated, and the testing of the Nonsense Words is initiated following the same procedure and sequence.

Subtest III: Oral Reading

The primary objective of the Oral Reading Subtest is to provide a measure of a child's ability to read words in context that have already been viewed in isolation (on the previous two subtests). The secondary objective is to provide an overall measure of the child's fluency in reading. The construction of Subtest III parallels that of Subtest I in that it is divided into the same eleven reader levels (covering Grades One through Five). One passage represents each of the eleven reader levels. Words are taken from the Subtest I reader-level lists and placed in passages at the same level designations. These words are labeled "Basal Target Words." Real Words are also

taken from the Phonic Decoding Subtest and placed in the passages. These words are labeled "Phonic Target Words." All of the other words in each passage are taken from reader levels below that of the Basal Target Words in the passage.

Construction of developmental version. Each passage is intended to represent a passage found in a reading series at a particular reader level. Therefore, three of the basal series used to develop Subtest I were analyzed for a variety of variables affecting the construction of reader passages (e.g., number of words per page, number of sentences per page) and a set of guidelines for writing the Oral Reading passages at each reader level were devised.

In order to fill the primary objective of Subtest III (that of measuring the contextual reading of the Target Words), the passages had to be constructed in such a way that all of the other words in each passage had a high likelihood of being "easier" for the child to read than the Target Words. Therefore, in constructing each passage, all non-target words were checked against the Harris & Jacobson list for grade level designations. Only words that were designated as below the grade levels of the Basal Target Words were used in a given passage as non-target words (with the obvious exception of the preprimer passage). From each list of 15 words, 10 words were selected on the basis of convienience in writing the passages.[4] The 10 words from each list then became the Basal Target Words in the passage at the corresponding reader level. A sequence of the Subtest II, Real Words was devised in which these words became "candidates" for the Subtest III passages by virtue of their Harris & Jacobson grade level designations. The general rule for the selection of the

Phonic Target Words was that they be "above grade level" with respect to the passage in which they were included. Five Phonic Target Words were selected for each of the Subtest III passages with the exception of the preprimer passage which includes no Phonic Target Words.

Response recording and administration procedures. The method for recording correct and incorrect responses to the Subtest III Target Words was the same as for Subtests I and II except for two modifications: (a) the "correct in over 5 seconds" category was eliminated, and (b) an incorrect response category of "meaningful substitution" was added, since it was considered that responses of this type (e.g., synonym substitutions, substitutions of one word for another word of the same speech part) might be useful in an evaluation of contextual reading ability.

In order to measure fluency in reading, a system of recording oral reading errors was devised which includes six different error types (referred to as Passage Errors): (a) Punctuation Error (obvious deletion or addition of a punctuation mark, primarily periods), (b) Error Substitution (replacement of a word in the passage with an incorrect word or a nonsense word), (c) Error Omission (failure to read a particular word), (d) Meaningful Substitution (replacement of a word in the passage with another word that preserves some meaning in the sentence. For example, when the word given by the child is a synoym for, or the same part of speech as the word replaced), (e) Word Addition (insertion of a word into the passage) and, (f) Transformation (alteration of a word in the passage in such a way that the root word remains basically correct. For example, addition or deletion of a word ending, such as *s* or *ing*; change in tense of a verb, such as *see* for *saw*). Two additional indices of oral reading fluency were obtained by requiring

the tester to evaluate the appropriateness of the child's "oral inflection" on a scale of one to three, and by measuring the amount of time required by the child to read each passage.

During test administration, the child is presented with a printed version of the passage, while the tester records all errors on a copy of the test protocol. An example of the tester's copy of the passage for the 3-1 reader level is shown in Figure 2. It will be noted from Figure 2 that each of the Basal Target Words in the passage is marked by a single underline, while each of the Phonic Target Words is marked by a double underline. At the bottom of the page, space is provided for scoring the target words, for tallying the oral reading errors, and for recording the Oral Inflection judgement and the reading time. During testing, the child's errors are recorded above the text of the passage. Following the session, this information is transfered to the spaces at the bottom of the page.

Subtest III is begun on the passage which corresponds to the child's 80% basal level list on Subtest I. If the child is correct on at least 8 (80%) of the Basal Target Words in that passage, the subtest is continued with the next passage until the child misses more than 5 (50%) of the Basal Target Words in a passage and the test is terminated. If the child is incorrect on more than 80% of the Basal Target Words on the first passage presented, the preceeding passage is given until the child is able to read at least 80% of the Basal Target Words. Then the successive passages are administered as described above.

Comparison of Good and Poor Readers

Subject Sample

Children were selected from the enire population of a Manhattan public school (Approximate N=600). The measures used for selection of the subject sample at each grade level are shown in Table 4. Using these measures, the 10 best and 10 poorest readers were selected at each of the six grade levels. The "poor" readers selected were from 1 to 3 years behind their grade-expected reading levels on the basis of all available test information. The "good" readers had obtained scores from "on grade level" to 2 years ahead of grade level on all available measures.

Table 5 shows the means ($\bar{x}$) and standard deviations (SD) of each subgroup of good (N=10) and poor (N=10) readers in each grade who received the two standardized group tests (Iowa Test of Basis Skills [ITBS] and New York City-Wide [NYC] Reading Test), an individually administered diagnostic test (Gilmore Oral Reading Test), and teachers' reports of the Basal Reader Placement levels of individual children. Reference to the ITBS data (for Grades Two through Four) reveals that the good readers averaged from 1½ to 2½ years (in grade equivalents) ahead of the poor readers. On the NYC scores (for Third through Sixth Graders), mean differences between the good and poor readers increases from approximately 2 years at Third Grade to nearly 3½ years at Sixth Grade. Even larger differences can be observed between good and poor readers on the Gilmore Oral. Finally, the Basal Reader Placements indicate differences of about 1 year between good and poor readers in the First through Third Grades.

Test Administration

Two days of training in the use of the test was conducted for a group of five testers. The training included studying the Manual

of Instructions, setting up mock testing situations, practicing response "coding" and "time training" for correct response latency estimates. Although the "time training" did not result in 100% accuracy in the testers' estimation of elapsed time, a fairly high degree of reliability (about 80%) was established.

Testing took place during the week of February 27 to March 3, 1978. On the first day of testing, when inter-tester reliability might be expected to be lowest, estimates of inter-tester reliability were obtained by conducting sessions in which one tester observed and completed a test protocol, while the second tester administered and also completed a protocol. A total of 14 such sessions were completed. Time required for complete test administration ranged from 20 minutes to 1 hour.

Results

Estimates of the test's reliability. As previously noted, in order to assess inter-tester reliability in the first field study, tester-pairs filled out matched protocols during 14 of the initial test sessions. The percentages of agreement on the coded response classifications (including latency estimates and incorrect response classifications) and on the simple classification of responses as correct and incorrect were computed for each list, subsection or passage. The results of this analysis are shown in Table 6. It will be noted from the table that the numbers of paired observations (N) varied from list to list. This is because no lists were used in the analysis that were above the ceiling or below the basal level for a given child. (In such cases, the percentage of agreement would automatically be 100.) It will be noted that the percentages of agreement on the specific response codes range from 72% to 100%,

with most above 85%.

Inter-tester agreement on the Right/Wrong classification (in which responses were considered independently of specific response codes), which has formed the basis for most of our analyses of these data, was much higher (ranging from 84% to 100%, with most above 90%). Our analysis of the responses on Subtest II (not shown in the table) yielded similar results.

Regarding internal consistency, the alpha coefficients generated on every item in the DST were all extremely high, ranging from .95 to .99. Although the issue of the test's reliability requires further investigation, it would seem from these preliminary analyses that it is an extremely reliable instrument.

Estimates of the test's validity. Estimates of the concurrent and convergent validity of the DST were made by computing correlations between the various DST scores and the academic measures available on our subject sample. Table 7 shows the correlations of the total correct on the Basal Word Recognition Subtest and the Real Word and Nonsense Word sections of the Phonic Decoding Subtest with these measures. Reference to the table reveals that all coefficients were high (and all were significant, $p<.001$). Particularly interesting is the pattern of these relationships. For example, the correlations with ITBS Reading (a comprehension measure) were lower than the corresponding correlations with ITBS Vocabulary (a word recognition measure). This same pattern may be observed in the correlations with the NYC scores, as well as with the Gilmore Oral scores. Moreover, the correlations with the two ITBS reading measures (Reading and Vocabulary) were both higher than those with the non-reading measures (Spelling and Math). This finding suggests a

high degree of convergent validity.

One additional observation should be made concerning these data. The correlations between the DST scores and the Program Placement information supplied by the teachers were higher than those with any other measure. Although this may, in part, be due to the fact that these placement measures were gathered only a few weeks before the DST was administered, it may also be due to the criterion-referenced nature of the DST. Research now in progress with larger samples will provide further evidence of the convergent and concurrent validity of the instrument.

Good and poor reader differences. Results of the 2-way analysis of variance (by Grade and Group) of the various DST scores are presented in Table 8. Five of the eight scores covered in this analysis are simple raw scores (i.e., total number of correct responses) on each of the subtest sections (BWR, PDRW, PDNW, BTW and PTW). Inspection of the table reveals that there were large and significant ($p<.001$) effects of both Grade and Group on all five of these measures. These results substantiate the fact that the DST reliably discriminates between good and poor readers, as well as between children at the grade levels tested (One through Six). Figures 3 and 4 show the trends of mean scores of the good and poor readers at each of the grade levels for three of these scores (BWR in Figure 3; PDRW and PDNW in Figure 4). The Grade x Group interaction terms (see Table 8) were significant for three of the five raw scores (PDRW, BTW and PTW) and approached significance in the other two cases. This interaction effect is a function of test ceiling. Many of the Fourth Fifth and Sixth Grade good readers "topped-out" on these measures (i.e., scored close to 100% correct), while the poor readers continued to display increasingly higher scores

from one grade to the next, without approaching the subtest (or subtest section) ceilings. This trend is particularly apparent in the Phonic Decoding-Real Word (PDRW) scores shown in Figure 4. The mean score of the Third Grade good readers on this measure was 55 correct, leaving little room for improvement by the good reader samples in successive grades (since the section contains only 60 items). On the other hand, the mean score of the Third Grade poor readers on the PDRW measure was only 24. Thus, the scores of the poor readers in the successive grade levels show improvement, to a high of 51 in the Sixth Grade sample. Where the interaction term is not significant, as for the Basal Word Recognition scores (see Table 8), we see that the gap between mean scores of good and poor readers on this measure (see Figure 3) remains fairly constant (around 50 points) across successive grade levels, and the good reader sample approaches the subtest ceiling only at the Fifth and Sixth Grade levels.

One other point to be raised in discussing the trend of mean raw scores of good and poor readers relates to the two Phonic Decoding scores shown in Figure 4. As can be seen in the figure, the mean scores of good readers (solid lines) on both the Real Word (PDRW) and Nonsense Word (PDNW) sections remain fairly similar at each of the grade levels, and eventually meet at the subtest ceiling in Grade Six. However, in the poor reader sample (broken lines) a different trend is observable. Mean scores in the First Grade poor reader sample on both PDRW and PDNW hover around zero and progressively grow farther apart. While scores on PDRW approach the ceiling in the Sixth Grade poor reader sample, those on PDNW show little progress between the Fourth and Sixth Grade samples (only

about 5 points) and remain far short of the sections's ceiling (28 points below) in the Sixth Grade. This trend indicates that, throughout their vocabulary development, good readers are able to read most nonsense (i.e., "unknown") words that correspond, in terms of phonetic structure, to phonetically regular real words which they can also read. On the other hand, this ability does not readily develop in poor readers, who, by the Sixth Grade, are able to read only about 50% of the corresponding nonsense words. This phenomenon is represented more clearly by our Phonic Transfer Index, which is discussed below.

The Phonic Transfer Index (PTI) was designed to represent the degree to which a child's ability to read real words that conform to common spelling patterns (i.e., the items on the Real Word section of the Phonic Decoding Subtest) transfers to the decoding of nonsense words which follow the same spelling patterns (i.e., items on the Nonsense Word section). This index exploits the parallel-item feature of Subtest II and is defined as follows:

$$PTI = \frac{PDNW/PDRW}{PDRW\text{-}Tot.}$$

Where PDNW/PDRW is the number of Nonsense Word items which were correct given that the corresponding Real Word items were correct; and PDRW-Tot. is the total number (out of 60) of Real Word items which were correct. A crucial element in the use of the PTI is that it reflects a child's ability to transfer the "phonic information" involved in the decoding of the Real Words to the decoding of the Nonsense Words, while it hold constant the actual level of mastery of the phonic patterns. For example, a child might be able to read only the short vowel (CVC and CCVCC) items on the Real Word section,

but be able to decode all of the corresponding Nonsense Word items. This would indicate a high level of phonic transfer and result in a high PTI. On the other hand, a child might be able to read most of the items on the Real Word section, but be unable to sound out more than a few of the corresponding Nonsense Word items. This would result in a low PTI. A PTI of 1.0 would indicate 100% transfer. That is, the child was able to read all of the Nonsense Word items that correspond to the Real Word items which he/she read correctly.

The analysis of the PTI shown in Table 8 reveals large and significant effects of both Grade and Group. It is intereseting to note that on this score the Group variable produces a much larger F-value (60.91) than does the Grade variable (4.85). This suggests that a child's achievement level relative to his age group (i.e., whether he is a good or poor reader in any given grade) is more important than grade level in determining the child's phonic transfer ability. Futhermore, the lack of significant (Grade x Group) interaction in the PTI analysis suggests that this difference between good and poor readers appears at the early grade levels and remains constant throughout the elemenatry school grades.

These suggestions are reinforced by the trend of mean PTIs of good and poor readers across the six grade levels, which is shown in Figure 5. Whereas the mean PTIs of the good readers in the sample show a steady increase across successive grade levels, from an approximate .57 in the First Grade to .99 in the Sixth Grade; the mean PTIs of poor readers show an erratic trend, with a low of .33 in the First Grade, to a high of .60 in the Second Grade, and a leveling off at approximately .55 in the Fifth and Sixth Grades.[5]

Thus, by the Sixth Grade, good readers are able to apply "phonic information" in decoding nonsense words at close to 100% efficiency, while poor readers can perform this function at only about 55% efficiency.

One further point concerning the strength of the Group variable in the PTI analysis (see Table 8) can be made by comparing mean scores abstracted from Figures 3, 4, and 5. Figure 3 shows that, on the BWR measure, mean scores of Second Grade good readers and those of Sixth Grade poor readers are very similar (117 and 122, respectively). In Figure 4, we see that this same phenomenon holds: Mean scores of Second Grade good readers and those of Sixth grade poor readers are again quite close (50 and 51, respectively). In other words, on the basis of the raw scores on both of these "real word" measures, we cannot distinguish a Second Grade good reader from a Sixth Grade poor reader. However, on the PDNW measure we see a much larger difference between the two groups ($\bar{x}$=45 for good readers; $\bar{x}$=32 for poor readers). This difference in performance on the Nonsense Word items results in widely different PTIs in these two groups: In terms of the phonic transfer skill, the Second Grade good reader is applying knowledge of phonic patterns to nonsense words at an approximate 80% efficiency; while the Sixth Grade poor reader is performing this function at only 50% efficiency.

The Passage Error Rate (PER) is the number of Passage Errors per 100 words of text in the Subtest III passages. Since there would be an obvious difference in the PERs of good and poor readers that would be due tc differing ability levels (at a given level of difficulty, poor readers simply will not recognize as many words as good readers), a technique was used that seeks to hold ability

level (in terms of word recognition) constant across the good and poor reader groups. Since the Target Words contained in a particular passage are presumed to be the most difficult words in the passage, a fairly high degree of word recognition ability would be assumed in a passage on which a child is able to read all or most of the Target Words. Thus, the "Basal Level" passage, defined as the highest passage on which each child is able to read at least 80% of the Basal Target Words, was selected for the analyses of PER[6] (see Table 8 and Figure 6). Therefore, in these analyses, we might be comparing the PER of a Third Grade poor reader on the 2-1 passage, with the PER of a Third Grade good reader on the 4-1 passage. Since the level of word recognition ability of the poor reader on the 2-1 passage would be approximately equal to that of the good reader on the 4-1 passage (i.e., 80% of the Basal Target Words), we would assume that differences in the PERs of these two children would not be exclusively the function of differing levels of word recognition ability.

The analysis of the PER shown in Table 8 indicates significant effects of both Grade and Group, as well as Grade x Group interaction. The Grade effect suggests that overall fluency in oral reading (as measured by the six types of Passage Errors included in Subtest III) changes as children grow older. In fact, as Figure 6 shows, it improves. Figure 6 also shows that the interaction effect shown in Table 8 is caused by the fact that the mean PERs of good and poor readers in the sample were extremely similar in the Third and Fourth Grades; while in the Second, Fifth and Sixth Grades, they were quite different. Future field studies involving the DST will verify whether this phenomenon is a function of the instrument or

of the specific sample involved in this field study. However, the significant effect of Group on this measure (which is clearly apparent in the differing mean scores of the good and poor readers in Grades Two, Five and Six) does suggest that fluency in oral reading differs between good and poor readers, and that this is the case even when we consider material that is within the range of recognizable vocabulary for poor readers.

Conclusions

It seems from this preliminary study that the DST is a highly reliable and valid instrument for the measurement of decoding skills. Although we have just begun to explore its reliability in this report, the results are promising. The degree of convergent validity, evidenced in the trend toward higher correlation coefficients with other measures of reading, as compared to measures of spelling and math, has not often been found for other individual reading tests (see McCall & McCall, 1969). Furthermore, the convergence of the DST scores on word recognition measures, as compared to comprehension measures, suggests that it may provide a more valid measure of decoding in reading than most previously available diagnostic reading tests. Finally, the fact that the DST measures correlate more highly with program placement than with test scores supports the validity of the criterion-referenced construct. Reading program placement is more closely associated with achievement in the curriculum sequence than with achievement on standardized tests. Thus, the high correlations between the DST scores and program placement indicate that the DST is accurately reflecting a child's position in the curriculum sequence.

The degree to which the test discriminates between the good and poor reader samples is remarkable. Our comparison of these two

groups indicates that the DST reliably meets the objective of identifying children who are having difficulty in reading. Without exception, there are large differences between the mean scores of the two groups across all grade levels. (The ranges of the five raw scores of the two groups [reported in Richardson, DiBenedetto & Adler, Note 1] do not even overlap until the Third Grade level).

The finding regarding transfer ability is particularly interesting. It will be recalled that, while the Second Grade Good readers resembled the Sixth Grade poor readers on both the Basal Word Recognition and Phonic Decoding/Real Words raw scores, they scored much higher than the Sixth Grade poor readers on the PTI. Our findings indicate that, while poor readers tend to lag behind good readers at each grade level on all measures of decoding, a failure to transfer letter-sound knowledge is a particularly distinguishing characteristic of poor readers, regardless of age or grade level. This might be a contributing cause, as opposed to a symptom, of reading problems. It is possible that good readers, who rapidly develop a facility with the use of letter sounds in decoding, use this ability to foster their own reading achievement. In contrast, poor readers, who seem to reach a limit on this ability far below its optimum, are unable to support their achievement and, consequently fall further and further behind in reading. At present our results must be considered preliminary in this regard. However, they do suggest that the investigation of transfer in reading disabled children may be an important topic for future research.

Reference Note

1. Richardson, E., DiBenedetto, B., & Adler, A A proposal for the extension of a project to develop a test of decoding skills. Submitted to National Instititute of Child Health and Human Development, Washington D.C., September, 1978 (Contract No. 1-HD-7-2837).

References

Bruininks, R.H., Lucker, W.G., & Gropper, R.L. Psycholinguistic abilities of good and poor reading disadvantaged first-graders. Elementary School Journal, 1970, 70, 378-386.

Botel, M. Botel predicting readability levels. Chicago: Follett, 1962.

Chall, J. Learning to read: The great debate. New York: McGraw-Hill, 1967.

Cohen, S.A. Studies in visual perception and reading in disadvantaged children. Journal of Learning Disabilities, 1974, 7, 429-35.

Frostig, M., Lefever, D.W., & Whittlesey. The Marianne Frostig development test of visual perception. Palo Alto, California: Consulting Psychological Press, 1966.

Golden, N.E., & Steiner, S. Auditory and visual functions in good and poor readers. Journal of Learning Disabilities, 1969, 2, 476-481.

Goldman, M., Fristoe, M., & Woodcock, R.W. Diagnostic auditory discrimination test. Circle Pines, Minnesota: American Guidance Service, 1974.

Hammill, D.D. & Larsen, S.C. The relationship of selected auditory skills and reading ability. Journal of Learning Disabilities, 1974, 7, 429-435.

Harris, A.J. & Jacobson, M.D. Basic elementary reading vocabularies. New York: MacMillan, 1972.

Kirk, S.A., McCarthy, J.J., & Kirk, W.D. Illinois test of psycholinguistic abilities (Rev. ed.). Urbana, Illinois: University of Illinois Press, 1968.

Larsen, S.C., Rogers, D., & Sowell, V. The use of selected perceptual tests in differentiating between normal and learning disabled

children. Journal of Learning Disabilities, 1976, 9, 32-37.

McCall, R.A., & McCall, R.B. Comparative vailidity of five reading diagnostic tests. Journal of Educational Research, 1969, 62, 329-33.

Macione, J.R. Psycholinguistic correlates of reading disability as defined by the Illinois test of psycholinguistic abilities. (Doctoral Dissertation, University of South Dakota, 1969). University Microfilms No. 70-5308.

Department of Health, Education, and Welfare. Request for Proposal (RFP# 77-18). Washington, D.C.: January 21, 1977.

Vellutino, F.R., Smith, H., Steger, J., & Kaman, M. Reading disability: Age differences and the perceptual deficit hypothesis. Child Development, 1975, 46, 487-493.

Venezky, R.L. The structure of English orthography. The Hague, Netherlands: Moulton, 1970.

Footnotes

[1]Subtest I has been subsequently revised. It now consists of 11 lists of 10 words each.

[2]For a comprehensive assessment of a child's knowledge of spelling-to-sound patterns in English, the DST system may be regarded as somewhat incomplete. For example, it fails to include such relationships as alternate consonant sounds, common word endings, and a number of other common linguistic markers and relational units.

[3]In the revised version of Subtest II, the monosyllabic words were selected to represent the Second through Sixth Grade levels.

[4]In the revised version of the test, all 10 words presented in a given list in Subtest I were included in the corresponding Subtest III passage.

[5]The mean PTI of .60 in the Second Grade poor reader group may be somewhat misleading. Since these children generally scored so low on PDRW (a mean score of 9.7, see Figure 4) only a small number of correct corresponding PDNW items would yeild a relatively high PTI.

[6]Errors on the Basal and Phonic Target Words were not included as Passage Errors. Children who had "topped out" (i.e., correctly read more than 80% of the Basal Target Words on all passages up through 5-2), as well as children who did not correctly read at least 80% of the Basal Target Words in the PP passage, were not included in these analyses.

Table 1

Reader Series Used in Preparation of Word Lists

Reading Program	Reader Level Word Lists Used						
	pp	p	1	2	3	4	5
Houghton-Mifflin (revised)	x	x	x	x	x	x	x
Macmillar "r"	x	x	x	x	x	x	
Book Mark	x	x	x	x	x	x	x
Allyn-Bacon (Sheldon Centennial)	x	x	x	x	x	x	x
Bank Street	x	x	x	x	x		
Macmillan (unrevised)	x	x	x	x	x	x	x
SRA/BRS	x	x	x	x			
Merrill Linguistic	x	x	x	x			
Programmed Reading (Sullivan)	x	x	x	x			
Lippincott		x	x	x	x	x	

Table 2

Basal Word Recognition Subtest for 2-1 Reader Level

Word	Frequency Information Mean.	Code	Prior Intro.	Cum. Tot.	Reader Level Value	Validation Information Botel	H-J
better	5	2	2	9	1.0	2-1	1
begin	5	3	0	8	1.2	3-1	2
dear	4	3	1	8	1.4	2-2	2
dry	4	2	3	9	1.2	3-1	2
end	5	4	0	9	1.1	2-2	2
fine	5	3	1	9	1.1	2-1	2
grew	5	2	2	9	1.4	2-2	2
kind	5	3	1	9	1.2	2-2	1
paw	5	2	2	9	1.4	3-1	2
part	5	3	1	9	1.2	2-2	2
river	5	3	1	9	1.2	2-2	2
shop	4	3	3	10	1.3	3-1	2
start	4	3	2	9	1.4	2-1	2
sound	5	3	2	10	1.4	2-2	1
wood	4	2	3	9	1.3	2-1	2

Table 3

Phonic Decoding Items (Subtest II)

	Monosyllabic Words			
	Single Consonants		Consonant Blends and Digraphs	
	Real Words	Nonsense Words	Real Words	Nonsense Words
Short Vowels	cat	lat	drop	brop
	hit	jit	shut	thut
	fed	med	path	sath
	nut	dut	stick	glick
	jog	nog	flesh	blesh
Long Vowels	game	bame	white	stite
	hide	mide	chase	drase
	cute	fute	drove	chove
	doze	voze	brake	grake
	fake	pake	grave	clave
Dipthtongs and Vowel Digraphs	town	cown	brown	strown
	loud	roud	threw	prew
	join	zoin	grain	thrain
	raw	taw	choice	froice
	loaf	woaf	bound	spound

	Polysyllabic Words			
Short Vowels	visit	lisit	finish	brinish
	pencil	jencil	pumpkin	thumpkin
	sunset	munset	publish	sublish
	bullet	dullet	splendid	glendid
	atomic	anomic	stiffen	bliffen
Long Vowels	beside	semide	telephone	stelephone
	mistake	fistake	grateful	clateful
	umpire	vumpire	provide	drovide
	rotate	motate	explode	excrode
	awoke	aboke	demonstrate	semonstrate
Diphthongs and Vowel Digraphs	cowboy	rowboy	flower	strower
	rooster	cooster	complain	comprain
	sausage	wausage	creature	threature
	avoid	atoid	embroider	emfroider
	announcer	azzouncer	applaud	aspaud

Table 4

Reading Measures Used In Subject Selection at Various Grade Levels

Reading Measures	Administration Period	Grade 1	2	3	4	5	6
New York City Reading Test	Mar. '77			x	x	x	x
Iowa Test of Basic Skills	May '77		x	x	x		
Phonic Skills Test	Oct. '77	x	x	x			
Gilmore Oral Reading Test	Nov. '77				x	x	x
Reading-Placement Information	Jan. '78	x	x	x			

Table 5

Scores of Good and Poor Readers at Various Grade Levels on Subject Selection Measures

Measure		First		Second		Third		Fourth		Fifth		Sixth	
		Good	Poor	Good	Poor	Good	Poor	Good	Poor	Good	Poor	Good	Poor
Iowa Test of Basic Skills[a] — Total Reading	$\bar{x}$	----	----	2.89	1.18	3.02	1.22	4.24	1.68	----	----	----	----
	SD	----	----	.30	.33	.25	.37	.47	.64	----	----	----	----
Iowa Test of Basic Skills[a] — National Percentile	$\bar{x}$	----	----	78.6	18.0	53.89	2.7	60.1	5.20	----	----	----	----
	SD			6.57	10.45	7.80	1.77	14.31	5.47	----	----	----	----
Iowa Test of Basic Skills[a] — Local Percentile	$\bar{x}$	----	----	87.8	17.1	81.89	10.20	90.8	16.7	----	----	----	----
	SD	----	----	6.43	9.86	6.31	7.27	5.92	12.76	----	----	----	----
City-Wide Test[b] — Reading Vocabulary	$\bar{x}$	----	----	----	----	3.33	1.44	4.89	2.44	6.24	2.92	6.71	3.35
	SD	----	----	----	----	.66	.45	.96	.28	1.11	.63	1.10	.69
City-Wide Test[b] — Reading Comprehension	$\bar{x}$	----	----	----	----	3.21	1.21	4.69	2.70	6.44	3.17	7.15	3.37
	SD	----	----	----	----	.33	.66	1.63	.58	1.80	.64	1.24	1.11
Gilmore Oral Reading[c] — Accuracy	$\bar{x}$	----	----	----	----	----	----	5.39	3.69	8.15	3.38	9.55	4.56
	SD	----	----	----	----	----	----	2.25	.98	1.51	.75	.41	.66
Gilmore Oral Reading[c] — Comprehension	$\bar{x}$	----	----	----	----	----	----	5.10	3.87	7.70	3.94	8.11	5.10
	SD	----	----	----	----	----	----	2.39	1.9	1.81	1.42	1.99	1.91
Basal Reader Placement	$\bar{x}$	1.8	1.1	2.6	1.8	3.5	2.3	----	----	----	----	----	----

[a]National norms for test administered in May, 1977 were (in Grade Equivalents): 1.9 (present second graders), 2.9 (present third graders), and 3.9 (present fourth graders).

[b]City-wide norms for test administered in March, 1977 (in Grade Equivalents): 2.7 (present third graders), 3.7 (present fourth graders), 4.7 (present fifth graders), and 5.7 (present sixth graders).

[c] Standardized norms for test administered in October, 1977 (in Grade Equivalents): 4.2 (fourth grade), 5.2 (fifth grade), and 6.2 (sixth grade).

Table 6

Percentage of Agreement on Response Classifications

Between Paired-Observations[a] on Subtests I and II

	Basal Word Recognition			Oral Passages		
List/Passage	N	Codes	Right/Wrong	N	Codes	Right/Wrong
pp	2	90%	98%	4	95%	98%
P	3	100	100	3	96	96
1	4	97	98	3	89	93
2-1	8	98	99	4	98	98
2-2	10	96	97	6	99	100
3-1	10	95	100	7	97	100
3-2	9	96	99	8	90	94
4-1	11	87	95	7	85	91
4-2	11	87	98	9	79	84
5-1	10	83	97	7	75	92
5-2	10	75	92	8	72	89

[a]Each paired-observation involved a total of 15 items. Thus, for example, the percentage of agreement computations on the PP lists are based on a total of 30 items for the two (2) paired-observations, and on the 2-1 lists, the percentages are based on a total of 120 items for the eight (8) paired-observations.

Table 7

Correlations of Basal Word Recognition and Phonic Decoding (Real Words and Nonsense Words) Raw Scores with Measures of Academic Achievement

Measures	N	Basal Words	Real Words	Nonsense Words
ITBS[a]:				
Reading	59	.78*	.75*	.79*
Vocabulary	59	.86*	.85*	.85*
Spelling	59	.69*	.67*	.63*
Math	57	.67*	.67*	.61*
NYC[b]:				
Vocabulary	77	.78*	.70*	.71*
Comprehension	77	.73*	.62*	.66*
Total Reading	77	.76*	.66*	.69*
Gilmore Oral:				
Accuracy	59	.78*	.68*	.71*
Comprehension	59	.53*	.44*	.38*
Program Placement[c]:				
Basal Reader	60	.92*	.89*	.83*
Phonics	60	.94*	.92*	.86*

[a]Iowa Test of Basic Skills

[b]New York City-Wide Reading Test

[c]Placement of children in reading programs in use in classroom. Supplied by teachers.

*$p < .001$

Table 8

F Values and Degrees of Freedom (df) Resulting from

2-Way ANOVA of Various DST Scores

Score	Grade		Group		Interaction		Error
	F-Value	df	F-Value	df	F-Value	df	df
BWR-Tot.[a]	59.42**	5	179.16**	1	2.20	5	108
PDRW-Tot.[b]	60.93**	5	193.04**	1	8.75**	5	108
PDNW-Tot.[c]	21.93**	5	160.79**	1	2.03	5	108
BTW-Tot.[d]	73.04**	5	227.30**	1	2.95*	5	108
PTW-Tot.[e]	57.64**	5	169.76**	1	5.28**	5	108
PTI [f]	4.85**	5	60.91**	1	.96	5	108
PER [g]	8.48**	5	16.87**	1	7.65**	5	55

[a]Basal Word Recognition-Total Raw Score

[b]Phonic Decoding Real Words-Raw Score

[c]Phonic Decoding Nonsense Words-Raw Score

[d]Basal Target Words-Raw Score

[e]Phonic Target Words-Raw Score

[f]Phonic Transfer Index

[g]Passage Error Rate

*$p<.05$

**$p<.001$

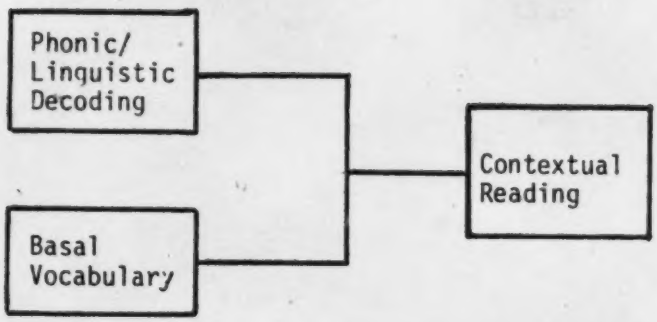

Figure 1. Simplified model of decoding in reading.

ORAL READING SUBTEST (Page 4)

THREE-ONE PASSAGE

One pleasant[1] day a man and his daughter[2] went out for a ride. Suddenly a pumpkin[1] bounced[3] from a fruit truck and hit their car. The man hit the brake.[2]

"Why are you dropping pumpkins on my car?" yelled the man.

A policeman walked up. "What's the problem?[4]" he asked.

"I want to complain.[3] This nut[4] is trying to take my pumpkins," said the truck driver.

"There must be some mistake.[5] I am a famous[5] star,[6]" said the man. "I don't want your pumpkins."

"Famous star indeed![7]" shouted the truck driver as he pushed a banana[8] into the man's face.

By this time there was a big crowd.[9] People began to throw fruit. Finally,[10] more policemen came and took the two men away. After the people had gone, only the man's poor daughter was left covered with fruit.

(Card #9 continued)

Basal Word Recognition				Phonic Decoding Words				Passage Scores		
Item	Word	Code	Col	Item	Word	Code	Col	Type	Score	Cols
1.	pleasant	____	38	1.	pumpkin	____	48	Time	___	53-55
2.	daughter	____	39	2.	brake	____	49	Punct	___	56-57
3.	bounced	____	40	3.	complain	____	50	Oral I	___	58-59
4.	problem	____	41	4.	nut	____	51	E. Sub	___	60-61
5.	famous	____	42	5.	mistake	____	52	E. Omis	___	62-63
6.	star	____	43					M. Sub	___	64-65
7.	indeed	____	44					Extra	___	66-67
8.	banana	____	45					Trans	___	68-69
9.	crowd	____	46							
10.	finally	____	47							

(End Card #9, Cols 70-79 blank, Col 80, 9)

Figure 2. Sample Test Protocol passage from Subtest III (Developmental Version).

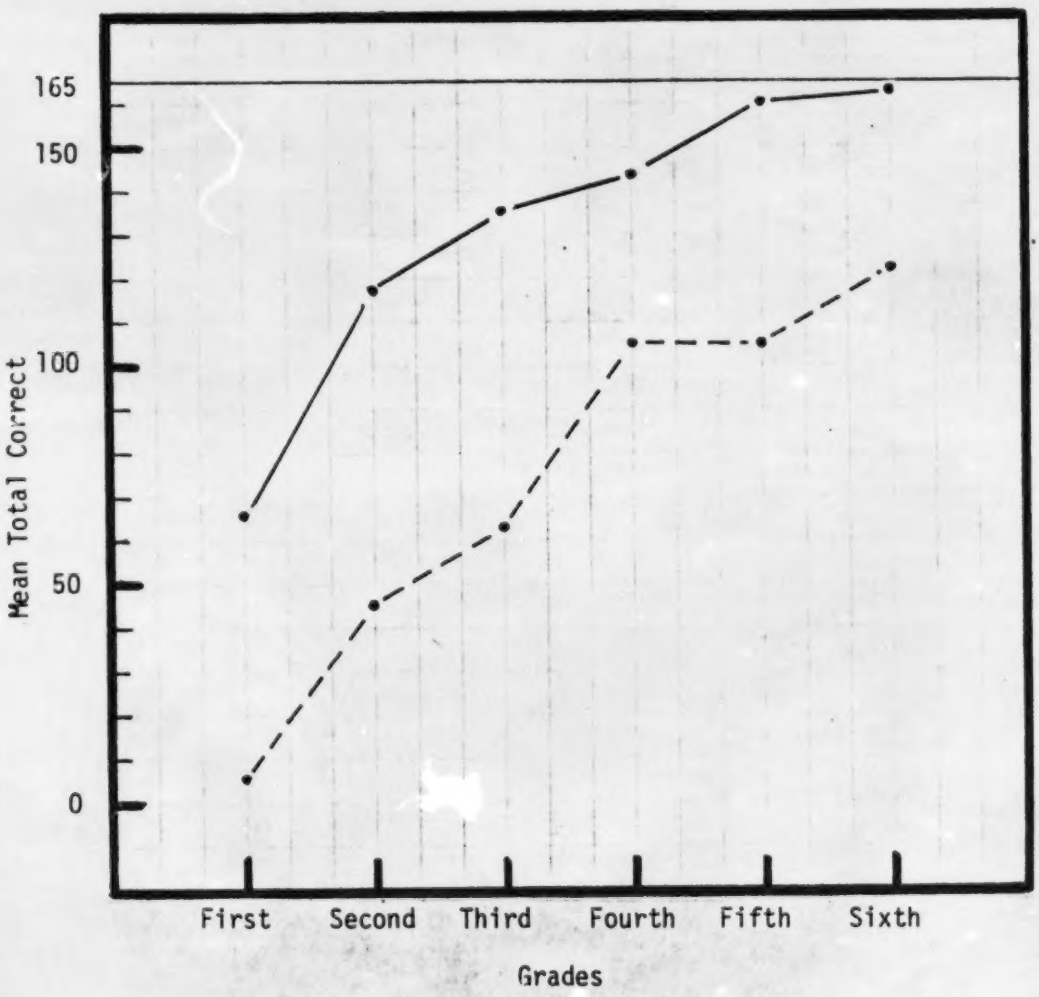

Figure 3. Mean total correct on Basal Word Recognition (BWR) Subtest for good reader (solid lines) and poor readers (broken lines) across grades.

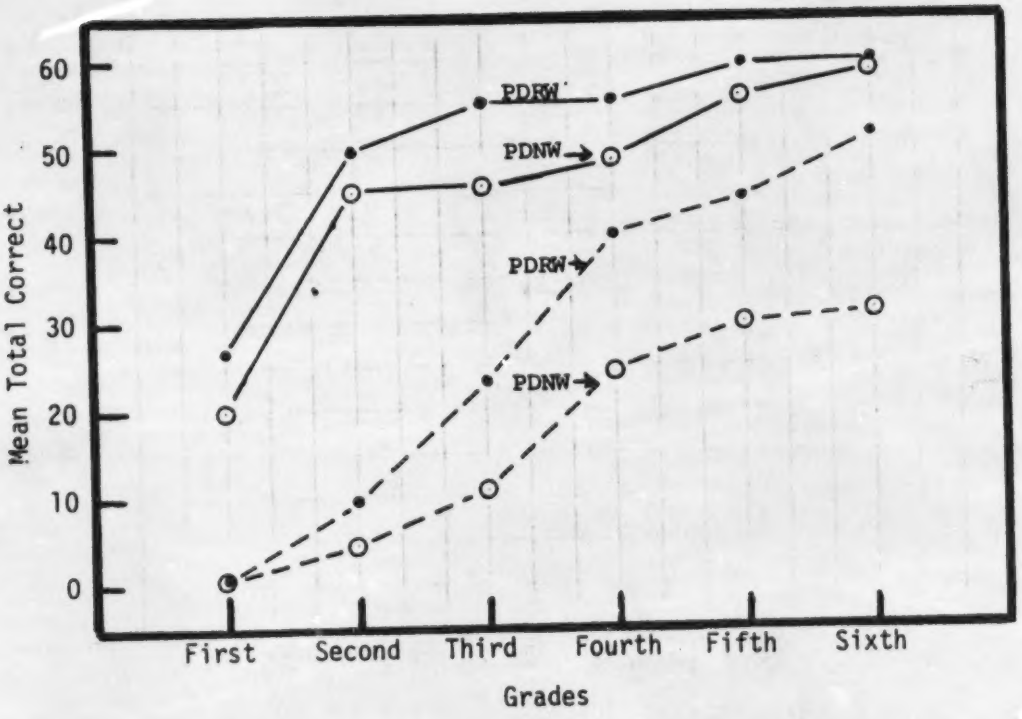

Figure 4. Mean total correct on Phonic Decoding, Real Word section (PDRW: closed points) and Nonsense Word section (PDNW: open points) for good readers (solid lines) and poor readers (broken lines) across grades.

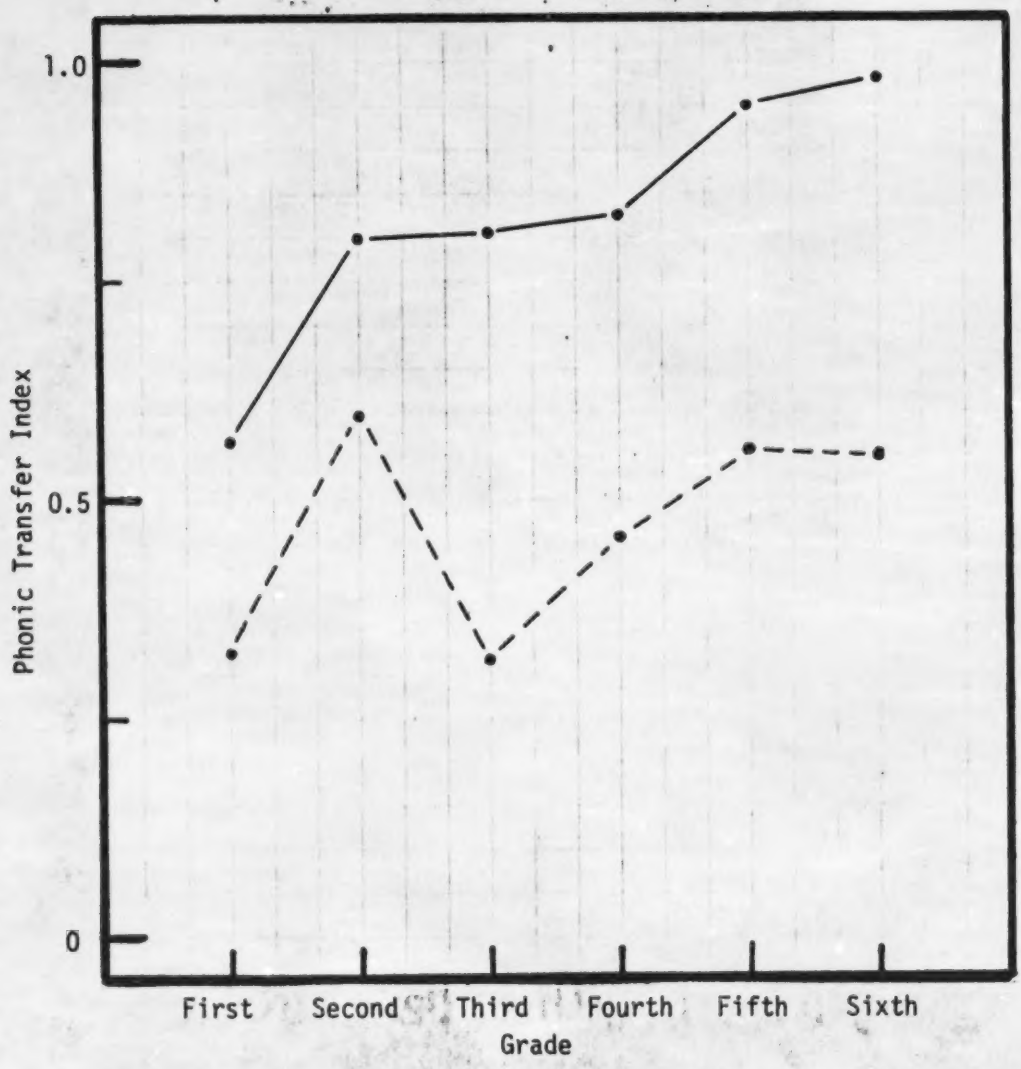

Figure 5. Mean Phonic Transfer Index (PTI) for good readers (solid lines) and poor readers (broken lines) across grades.

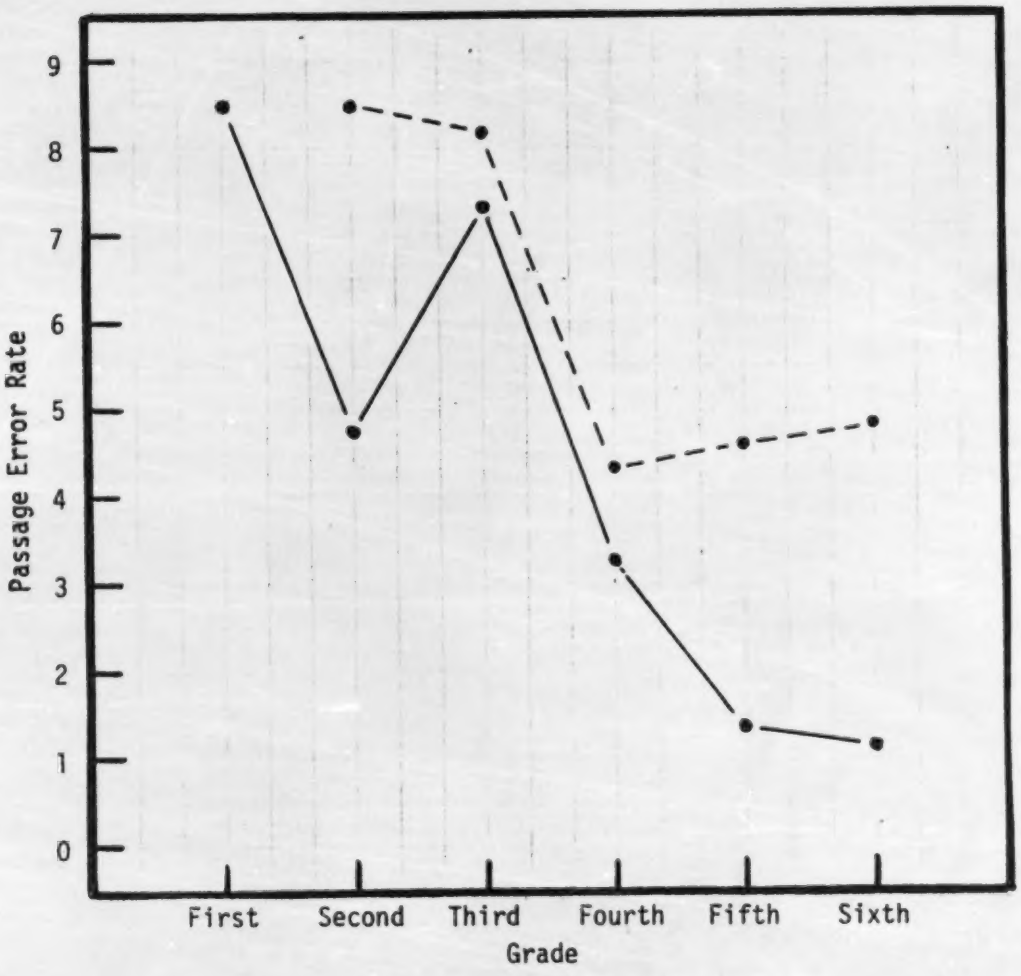

Figure 6. Passage Error Rate (PER) on Basal Level passages (Oral Reading Reading Subtest) of good readers (solid lines) and poor readers (broken lines) across grade levels.

Note. Children who did not obtain Basal Levels on Subtest III were not included in this analysis.